Chemistry & Biology, Vol. 11, 87–97, January, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2003.12.018

Biosynthesis of the Angiogenesis Inhibitor Borrelidin by *Streptomyces parvulus* Tü4055: Cluster Analysis and Assignment of Functions

Carlos Olano,[1] Barrie Wilkinson,[2] César Sánchez,[1] Steven J. Moss,[2] Rose Sheridan,[2] Vidya Math,[2] Alison J. Weston,[2] Alfredo F. Braña,[1] Christine J. Martin,[2] Markiyan Oliynyk,[3] Carmen Méndez,[1] Peter F. Leadlay,[2,3] and José A. Salas[1,*]

[1]Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (I.U.O.P.A.)
Universidad de Oviedo
33006 Oviedo
Spain

[2]Biotica Technology Ltd, Chesterford Research Park
Little Chesterford
Essex CB10 1XL

[3]Department of Biochemistry
University of Cambridge
80 Tennis Court Road
Cambridge CB2 1GA
United Kingdom

Summary

The biosynthetic gene cluster for the angiogenesis inhibitor borrelidin has been cloned from *Streptomyces parvulus* Tü4055. Sequence analysis indicates that the macrolide ring of borrelidin is formed by a modular polyketide synthase (PKS) (*borA1-A6*), a result that was confirmed by disruption of *borA3*. The borrelidin PKS is striking because only seven rather than the nine modules expected for a nonaketide product are encoded by *borA1-A6*. The starter unit of the PKS has been verified as *trans*-cyclopentane-1,2-dicarboxylic acid (*trans*-1,2-CPDA), and the genes involved in its biosynthesis identified. Other genes responsible for biosynthesis of the nitrile moiety, regulation, and self-resistance were also identified.

Introduction

Macrolide antibiotics are a large and structurally diverse class of natural products that includes compounds possessing potent and valuable therapeutic activities. Polyketide macrolides are synthesized through the repeated condensation of simple carboxylic acid units in a process closely resembling that of fatty acid biosynthesis. Their biosynthesis is catalyzed by polyfunctional type-I polyketide synthases (PKSs), also called modular PKSs, which are organized into repeating modules. Each module consists of the set of catalytic "domains" required to perform one round of chain assembly. Condensation of the next extender carboxylic acid unit onto the growing polyketide chain is performed by the catalytic activity of the essential β-ketosynthase (KS) domain. The acyl transferase (AT) and the acyl carrier protein (ACP) domains present in each module are responsible for the choice of extender unit and the retention of the growing chain and its transfer on the PKS, respectively. In addition, modules of type I PKSs may contain domains with ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) activities that are responsible for the reductive processing of the newly formed β-keto groups during chain extension [1, 2]. The final polyketide chain is released from the PKS by the action of a thioesterase (TE) domain, which is also generally involved in the cyclization of the final product [3]. In general, in a modular PKS there exists a correlation between the number of modules present in the PKS and the number of units incorporated. This correlation is termed colinearity. In addition, there is a single enzymatic domain present in the PKS for each successive catalytic step that occurs during biosynthesis, and they are used in a defined sequence. This mechanism is termed processive. This colinear and processive mechanism appears to be general for the biosynthesis of aliphatic polyketides by type I PKSs in bacterial systems [4].

Borrelidin (*bor*) is an 18-membered macrolide produced by several streptomycete species. The structure of borrelidin (Figure 1) was first elucidated in 1967 [5], then subsequently refined by detailed NMR analysis and confirmed by X-ray crystallography [6, 7]. The previously described antibiotic treponemycin was shown to be identical to borrelidin [8]. Borrelidin was discovered due to its antibacterial activity [9, 10] that involves selective inhibition of threonyl tRNA synthetase [11], and it also exhibits antiviral activity [12]. The recent discovery of antiangiogenesis activity highlights the potential of borrelidin as a therapeutic agent [13]. Angiogenesis is the formation of new vascular networks, and many inhibitors of the complex pathways involved in this process are being studied as potential targeted therapeutics for cancer chemotherapy. Tumor angiogenesis is provoked by the tumor responding to hypoxia, and the downstream consequences of this are mostly host-derived processes. Targeting tumor angiogenesis avoids the hurdles of other anticancer therapeutic modalities, such as the diversity of cancer types and drug resistance [14].

In addition to its potent angiogenesis-inhibiting activity, borrelidin was shown to possess antiproliferative activity toward various cell lines. Borrelidin inhibits tRNA synthetase and protein synthesis in cultured rat cells. However, the IC_{50} value for antiangiogenesis activity is 50-fold lower than that reported for inhibition of protein synthesis, indicating that these represent different activities of the compound [13]. Borrelidin has also been identified as an inhibitor of cyclin-dependant kinase Cdc28/Cln2 of *Saccharomyces cerivisiae*. It was found to arrest both haploid and diploid cells in late G_1 phase at concentrations that do not affect gross protein biosynthesis [15]. Present available data therefore indicate that borrelidin has potential as a lead compound to develop anti-tumor agents. This is reflected in the number of groups attempting synthetic programs toward borreli-

*Correspondence: jasalas@correo.uniovi.es

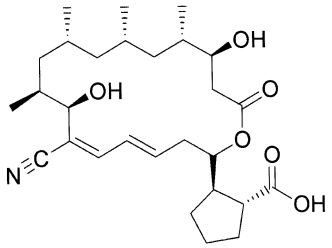

Figure 1. Structure of Borrelidin

din, with the first total synthesis reported recently by the Morken group [16].

Here, we report the cloning, sequencing, and characterization of the biosynthetic gene cluster for borrelidin in *Streptomyces parvulus* Tü4055. The cluster comprises genes encoding a type-I PKS, genes participating in the biosynthesis of the unusual starter unit *trans*-cyclopentane (1*R*,2*R*)-dicarboxylic acid (*trans*-1,2-CPDA), genes involved in further modification of the polyketide, and a borrelidin self-resistance gene. Insertional inactivation mutants were generated for every non-PKS gene within the biosynthetic gene cluster. Further, each mutant strain was complemented *in trans* by the addition of a full-length copy of the inactivated gene, and each was chemically complemented through the addition of exogenous *trans*-1,2-CPDA where appropriate. This leads to a detailed proposal for borrelidin biosynthesis with a functional assignment of the gene products from the cluster.

Results

Cloning and Sequencing of the Borrelidin Gene Cluster

A cosmid library of *S. parvulus* Tü4055 genomic DNA was screened by in situ colony hybridization using as a probe the 1.7 kb BglII-BamHI fragment of ORFB from *Streptomyces antibioticus* ATCC11891 [17]. This gene encodes the third subunit of the oleandomycin PKS and contains the coding region for the KS domain of module 6. Several positive clones were isolated and verified by Southern hybridization using the same probe as above. To test the possible involvement of these cosmids in borrelidin biosynthesis, two BamHI fragments (2.8 and 2.0 kb) from cosmid Bor32A2 were subcloned into the suicide vector pOJ260, and the resultant constructs (pOJM7 and pOJM8) were used for gene disruption experiments. These constructs were used to transform *S. parvulus* Tü4055 protoplasts, and the resultant apramycin-resistant strains were tested for their ability to produce borrelidin. HPLC analysis revealed both strains, designated as SPM7 and SPM8, respectively, to be deficient in borrelidin production. DNA sequencing later revealed that the two fragments used to generate the disruptants are adjacent to each other within the *bor* PKS. These experiments demonstrated that the cloned DNA region is involved in borrelidin biosynthesis. Several fragments of DNA that originate from the ends of the insert in cosmid Bor32A2 were used to screen the cosmid library. Ten new overlapping cosmid clones were identified, covering, in total, a chromosomal length of approximately 140 kb.

The complete nucleotide sequence of two of these overlapping cosmid clones, cosBor32A2 and cosBor19B9, was determined (EMBL accession number AJ580915). Analysis of the sequence revealed the presence of 42 ORFs. All of them showed the high G+C content and G+C bias in the third codon position that is characteristic of *Streptomyces* genes. Analysis of the deduced products of the different ORFs revealed the similarities shown in Table 1. Several ORFs (*borA1-borA6*) appear to encode for proteins that, based on the similarities to the conserved amino acid sequences of modular PKSs, clearly suggested their involvement in borrelidin biosynthesis. Some genes flanking the PKS genes were also assigned to be involved in borrelidin biosynthesis based on (1) comparison to proteins in databases and adscription of putative functions in borrelidin biosynthesis and (2) generation of mutants by gene disruption/replacement with the loss of the capability to produce borrelidin (see below). Some other ORFs (B1, B2, B11 to B17, B22; indicated in white in Figure 2) either did not show significant similarities to proteins in databases that would suggest a role for these proteins in borrelidin biosynthesis or their putative functions were found not to be necessary for borrelidin biosynthesis. In addition, several other ORFs (indicated in gray in Figure 2) showed significant similarities to proteins from *Streptomyces coelicolor* A3(2) (a borrelidin nonproducer), either showing the same genetic organization as in *S. coelicolor* A3(2) (B8 to B10) or showing a different genetic arrangement (B3 to B7 and B18 to B21). The disruption of *bor*B2 did not affect borrelidin production (data not shown). Based on these considerations, the ORFs indicated by the letter B (Figure 2) were tentatively excluded as being part of the borrelidin gene cluster, thus establishing the limits of the biosynthetic gene cluster.

Organization of the Borrelidin Gene Cluster

Genes proposed to encode enzymes involved in borrelidin biosynthesis span a region of approximately 52 kb. Six genes (*borA1-A6*) coding for the PKS are located in the middle of the cluster. Flanking the PKS genes on the left-hand side and transcribed in the opposite direction, are seven genes (*borB*, *borC*, *borD*, *borE*, *borF*, *borG* and *borH*). To the right-hand side of the PKS genes and transcribed in the same direction are six additional genes (*borI*, *borJ*, *borK*, *borL*, *borM*, *borN* and *borO*) (Figure 2). Some genes may be translationally coupled as they include overlapping stop and start codons (*borB* and *borC*, *borD* and *borE*, *borE* and *borF*, *borG* and *borH*, *borA3* and *borA4*, *borA4* and *borA5*, *borA5* and *borA6*, *borA6* and *borI*, *borJ* and *borK*).

Polyketide Synthase Genes

Six genes, *borA1* to *borA6*, encode a type I PKS containing a total of 31 individual domains organized into a loading module and six extender modules (Figure 3). All six genes are arranged head-to-tail and transcribed in the same direction. No apparent transcriptional terminator sequences were observed through DNA sequence analysis. Strikingly, sequence analysis of the DNA region

coding for the borrelidin PKS revealed several long repeated sequences (Figure 3). A region of 1001 bp was found to be repeated within *borA2* (module 1), *borA3* (module 2), and *borA6* (module 8). These repeated sequences comprise most of the coding regions for malonyl-specific AT domains and are identical in the three modules with the exception of 2 bp. Furthermore, modules 1 and 8 also share an additional DNA region of 168 bp of identity. Similarly, another DNA region of 803 bp is repeated within *borA3* (module 3), *borA4* (module 4), and *borA5* (module 5) with only one bp change (in module 5). Most of the coding regions for the KS and methylmalonyl-specific AT domains were included within these repeated sequences. Furthermore, modules 3 and 5 also share two identical regions of 560 and 304 bp located upstream and downstream, respectively, of the 803 bp region.

The *borA1* gene encodes the loading module. Assignment of the start codon is not obvious for *borA1*. There are four possible start codons prior to the beginning of the AT_0 domain sequence. The first start codon (position 16,184) leaves a significant N-terminal tail of 321 amino acids preceding the AT_0 domain. By comparison, the N-terminal tail preceding the AT_0 of the erythromycin PKS loading module is 108 amino acids, and that of the avermectin PKS loading module is 28 amino acids. It is therefore possible that one of the other candidate start codons could be correct (positions 16,298, 16,607, or 16,901). The length of the N-terminal tail suggests that it could possibly contain a catalytic activity, although it does not have any significant similarity to other sequences in databases. The BorA1 protein contains two domains encoding for an AT_0 and an ACP, respectively. The AT_0 domain of the borrelidin PKS loading module diverges from the majority of PKS AT domains, as the active site contains a cysteine residue instead of the usually observed serine such that the active site motif is GXCXG (specifically GHCYG). To date, for most type I PKS AT domain sequences, the conserved active site motif is GXSXG, which is the motif observed in lipases, fatty acid synthases, and most thioesterases.

BorA2 (one module), BorA3 (two modules), BorA4, BorA5, and BorA6 (all containing one module) encode the remainder of the borrelidin PKS. All of them contain conserved KS and ACP domains. The AT domains for these PKS extension modules display the active site motif GXSXG and also contain the expected motifs for the selection of either malonyl-CoA or methylmalonyl-CoA [18, 19]. The malonyl-CoA selective AT domains (AT1, AT2, and AT6) exhibit very high similarity to each other both at the protein and DNA levels. The same is true for the methylmalonyl-CoA-selective AT domains (AT3, AT4, and AT5); two of these AT domains (AT3 and AT4) have identical amino acid sequences throughout the conserved region. The high similarity of AT5 and AT4 to AT3 provides evidence that the extender unit selected by module 3 is methylmalonyl-CoA, and it is therefore most probable that the borrelidin C12-methyl group thus incorporated is subsequently modified to a nitrile moiety after incorporation into the PKS. All extension modules contain a KR domain, modules 2 and 3 include additional DH domains, and module 5 contains DH and ER domains. The final PKS gene, *borA6*, encodes the final extension unit, which incorporates malonyl-CoA and also contains the chain-terminating thioesterase domain (TE).

Curiously, although borrelidin is a nonaketide (which should require one loading plus eight extension steps), only seven modules are present in the cluster (one loading and six extension modules) as encoded by genes *borA1* to *borA6* (Figure 4). The possibility that DNA rearrangements could have occurred during construction of the cosmid library was excluded by extensive Southern hybridization analysis both in *S. parvulus* Tü4055 and in two other borrelidin producing streptomycetes (data not shown). Therefore, there are two condensation steps incorporating methylmalonate units for which no modules could initially be assigned. The implications of these findings are discussed below.

The gene *borB* encodes a thioesterase. BorB shows similarities to type II thioesterases such as PimI from *S. nataliensis* (49% identity; accession number CAC20922) and RifR from *Amycolatopsis mediterranei* S699, (48% identity; accession number AAG52991). Type II thioesterases have been proposed to rid the PKS of aberrant acyl chains that might otherwise block the enzyme complex [20, 21]. In the case of the tylosin type II thioesterase (Orf5), the in vitro behavior is consistent with its proposed role as an editing enzyme [22]. Disruption of type II thioesterases leads to a decreased level of polyketide production [20, 21]. Similarly, we found a 25% decrease in borrelidin production when *borB* was disrupted.

Genes Involved in the Biosynthesis of the Starter Unit

A retrobiosynthetic analysis of borrelidin suggests that *trans*-cyclopentane-(1*R*,2*R*)-dicarboxylic acid (*trans*-1,2-CPDA) may be the starter unit. This was confirmed by feeding the commercially available racemate of this acid to strains deficient in starter biosynthesis (see below). Several genes in the cluster located both upstream and downstream of the PKS genes are potential candidates to be involved in starter unit biosynthesis (see Table 1 for similarities). The *borN* gene product resembles an isomerase involved in the production of 2-oxohepta-3-ene-1,7-dioate, a key step in the catabolism of tyrosine via 4-hydroxyphenyl acetic acid (4HPA), which could be an intermediate in the biosynthetic pathway to *trans*-1,2-CPDA [23]. The *borE* product displays similarity to *O*-succinylbenzoyl-CoA synthase and chloromuconate cycloisomerases such as YkfB from *Bacillus subtilis* (26.2% identity; accession number AJ002571), which are enzymes belonging to the enolase superfamily. Members of the enolase superfamily share the ability to stabilize the formation of an anion on the carbon atom adjacent to a carboxylate group [24]. We therefore propose that BorE may be required for the cyclization of a putative intermediate in the biosynthesis of the starter unit. The *borC*, *borD*, *borK*, and *borM* genes could code for oxidoreductases resembling 2,5-dichloro-2,5-cyclohexadiene-1,4-diol dehydrogenase, 3-oxoacyl-ACP reductase, alcohol dehydrogenease Adh1, and F420-dependent dehydrogenase, respectively (Table 1). The products of three other genes, *borF, borG*, and *borH*, did not show significant similarities with proteins in data-

Table 1. Deduced Functions for Genes in the Borrelidin Gene Cluster

Gene	Amino Acids	Closest Similar Protein (% Identity/Similarity), Accession No.	Proposed Function in Borrelidin Biosynthesis
orfB1	103[a]	Putative integral membrane protein SCE22.09 from *S.coelicolor* (30/56), CAB90976	
orfB2	868	Hypothetical protein SCM2.07 from *S. coelicolor* (85/91), CAB65635	
orfB3	212	Hypothetical protein SCF76.07 from *S. coelicolor* (89/96), CAB56727	
orfB4	329	Probable AraC transcriptional regulator SCF76.06 from *S. coelicolor* (84/91), CAB56726	
orfB5	276	Nonheme chloroperoxidase (EC 1.11.1.10) SCF76.05c from *S. coelicolor* (89/94), CAB56725	
orfB6	185	Hypothetical protein SCF76.09 from *S. coelicolor* (76/84), CAB56729	
orfB7	323	Hypothetical protein SCF76.08c from *S. coelicolor* (80/88), CAB56728	
borB	264	Type II thioesterase PteH from *S. avermitilis* (52/68), AB070949	type II thioesterase
borC	265	2,5-dichloro-2,5-cyclohexadiene-1,4.diol dehydrogenase XF1726 from *Xylella fastidiosa* strain 9a5c (42/64), AAF84535	starter unit biosynthesis
borD	250	3-oxoacyl-ACP reductase FabG from *Plasmodium falciparum* (31/53), AF237573	starter unit biosynthesis
borE	390	O-succinylbenzoate-CoA synthase FN1586 from *Fusobacterium nucleatum* subsp. *nucleatum* (27/51), AAL93701	starter unit biosynthesis.
borF	272	Putative lysophospholipase homolog At1g73480 from *Arabidopsis thaliana* (26/49), AY045929	starter unit biosynthesis
borG	539	Acetolactate synthase, large subunit, MTH1444 from *Methanothermobacter thermautotrophicus* (30/52), AAB85919	starter unit biosynthesis
borH	675	Hypothetical protein PA3592 from *Pseudomonas aeruginosa* (31/54), AAG06980	starter unit biosynthesis
borA1	876	Polyketide synthase loading and extender module 1, SpnA from *Saccharopolyspora spinosa* (41/62), AAG23264	PKS loading domain (AT-ACP)
borA2	1571	Modular polyketide synthase PteA4 from *S. avermitilis* (54/72), BAB69306	PKS module 1(KS-ATa-KR-ACP)
borA3	3500	Modular polyketide synthase PteA2 from *S. avermitilis* (55/72), BAB69304	PKS modules 2 and 3 (KS-ATa-DH-KR-ACP-KS-ATp-DH-KR-ACP)
borA4	1620	Modular polyketide synthase OlmA7 from *S. avermitilis* (54/71), BAB69196	PKS module 4 (KS-ATp-KR-ACP)
borA5	2156	Soraphen polyketide synthase SorA from *Polyangium cellulosum* (51/69), AAK19883	PKS module 5 (KS-ATp-DH-ER-KR-ACP)
borA6	1742	Modular polyketide synthase PteA5 from *S. avermitilis* (50/68), BAB69307	PKS module 8 (KS-ATa-KR-ACP-TE)
borI	426	Cytochrome P450 TylH1 from *S. fradiae* (41/63), AAD12167	nitrile biosynthesis
borJ	454	DAPA aminotransferase BioA from *Kurthia* sp (39/62), BAB39453	nitrile biosynthesis
borK	330	Alcohol dehydrogenase Adh1 from *Aquifex aerolicusv* (36/56), AAC07327	unknown
borL	446	Putative auxin-regulated protein GH3 At1g59500 from *Arabidopsis thaliana* (24/49), NP_176159	starter unit biosynthesis
borM	305	Hypothetical protein SCL6.10 from *S. coelicolor* similar to putative F420-dependent dehydrogenase (36/54), CAB76875	starter unit biosynthesis
borN	248	Hypothetical protein SC1C2.27 from *S. coelicolor*, 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase superfamily (48/65),CAB55708	starter unit biosynthesis
borO	675	Protein ThrS from *Mycobacterium tuberculosis* H37Rv, similar to threonyl-tRNA synthetase (50/69), CAB08628	borrelidin resistance
orfB8	126	Hypothetical protein SC5E9.01c from *S. coelicolor* (81/86), CAC14481	
orfB9	127	Putative anti-sigma factor antagonist SC5E9.02 from *S. coelicolor* (54/68), CAC14482	
orfB10	82	Hypothetical protein SC5E9.03 from *S. coelicolor* (79/88), CAC14483	
orfB11	70	Hypothetical, no hits.	
orfB12	172	Putative regulator of the two componen systems SC5C11.03 from *S. coelicolor* (41/57), CAB76309	
orfB13	163	Putative acetyltransferase SCM2.08c from *S. coelicolor* (68/78), CAB65636	
orfB14	638	Putative secreted protein SC8A11.15c from *S. coelicolor* (51/67), CAC01587	
orfB15	165	Hypothetical protein SC9H11.25c from *S. coelicolor* (29/48), CAB92214	
orfB16	783	Putative oxidoreductase SCBAC25F8.17 from *S. coelicolor* (84/91), CAC42152	
orfB17	177	Conserved hypothetical protein SCBAC25F8.16 from *S. coelicolor* (58/73), CAC42151	
orfB18	251	Product unknown, Orf251 from *S. aureofaciens* (90/94), AAD23399	

(*continued*)

Table 1. Continued

Gene	Amino Acids	Closest Similar Protein (% Identity/Similarity), Accession No.	Proposed Function in Borrelidin Biosynthesis
orfB19	467	Putative aldehyde dehydrogenase AldA from *S. aureofaciens* (87/92), AAD23400	
orfB20	370	Putative aldehyde dehydrogenase SC9B5.09 from *S. coelicolor* (77/87), CAA22751	
orfB21	301	Hypothetical protein SC9B5.10 from *S. coelicolor* (71/80), CAA22752	
orfB22	180	Putative calcium binding protein SC10F4.20 from *S. coelicolor* (58/69), CAC16980	

[a] Incomplete ORF.

bases that can suggest a defined role in starter unit biosynthesis (Table 1). The *borL* product shows similarity to auxin response proteins from plants. Auxins are hormones involved in the regulation of various cellular processes in plants [25].

To investigate the role of these genes in *trans*-1,2-CPDA biosynthesis, individual mutants in all these genes were generated by insertional inactivation. They were inactivated in a manner designed to avoid the possibility of polar effects, which was verified in all cases by in *trans* complementation with a full-length copy of the disrupted gene under the control of the erythromycin resistance promoter *ermE**. Two groups of mutants were found: (1) mutants in which borrelidin production was abolished and (2) mutants in which borrelidin was produced, but the yields were clearly diminished (75%–89% decreased). The first group included mutants in *borC*, *borD*, *borE, borF, borK, borL*, and *borM*, and the second group included mutants in *borG, borH*, and *borN*. To further investigate the participation of these genes in *trans*-1,2-CPDA biosynthesis, feeding experiments were carried out. The mutants were grown in the presence of *trans*-1,2-CPDA, and borrelidin production was assessed (Table 2). Mutants in the first group (genes *borC*, *borD*, *borE, borF, borK, borL*, and *borM*) recovered the ability to synthesize borrelidin after *trans*-1,2-CPDA addition, and the levels of production were similar or even higher than in the wild-type strain. Mutants in the second group (genes *borG, borH*, and *borN*) clearly increased their levels of production, even exceeding, in most cases, the levels of the wild-type strain. Consequently, it can be deduced that genes *borC*, *borD*, *borE, borF, borK, borL*, and *borM* are essential for starter acid biosynthesis. The genes *borG, borH*, and *borN* may well be involved in the biosynthesis of *trans*-1,2-CPDA, but these data and protein similarity searches do not suggest a role. However, it may be that these genes have no role whatsoever in *trans*-1,2-CPDA biosynthesis, as feeding the wild-type strain with exogenous starter acid likewise increases the titer of borrelidin. The fact that the level of borrelidin production was increased by addition of *trans*-1,2-CPDA both in the different mutants and in the wild-type strain suggests that biosynthesis of *trans*-1,2-CPDA might be a bottleneck during borrelidin biosynthesis in this strain. Borrelidin production increased when additional copies of genes *borE* and *borL* were introduced into *S. parvulus* Tü4055. The genes were expressed in a multicopy plasmid vector under the control of the strong constitutive promoter *ermE**. Biosynthesis of borrelidin increased 4.2 $\pm$ 0.3 and 4.3 $\pm$ 0.7 fold upon expression of *borE* and *borL*, respectively, when compared with the titer of borrelidin of the wild-type strain only containing the vector.

Genes Involved in the Formation of the Nitrile Moiety

Borrelidin contains a nitrile moiety at C12 of the macrolide ring. Sequence analysis of the AT domain of the bor PKS module 3 in BorA3 indicates that the substrate utilized for the third round of chain extension is methylmalonyl-CoA (see above). Thus, the carbon atom of the nitrile moiety most probably arises from the methyl group of methylmalonyl-CoA. The products of the genes *borI* and *borJ* may be required for the formation of the nitrile moiety at C12. The *borI* gene product would encode a cytochrome P450 hydroxylase. BorI shares the highest similarities with TylHI, which catalyzes the hydroxylation of an exocyclic methyl group of the tylosin

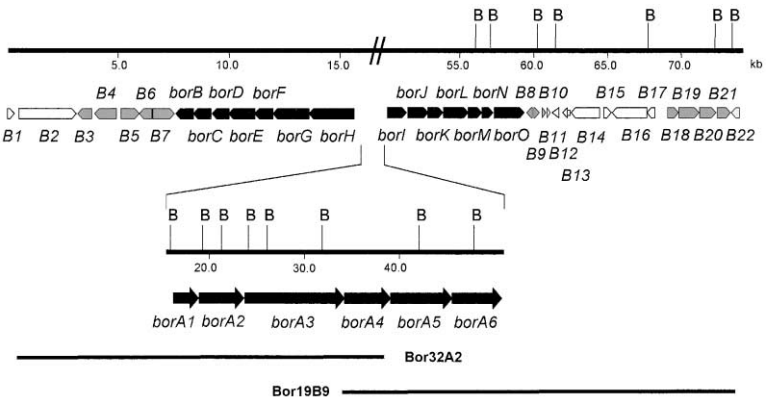

Figure 2. Schematic Representation of the Borrelidin Gene Cluster

B, BamHI. ORFs involved in borrelidin biosynthesis are colored in black and those not involved in gray or white, the gray ones corresponding to homologous genes in *S. coelicolor* showing the same genetic organization.

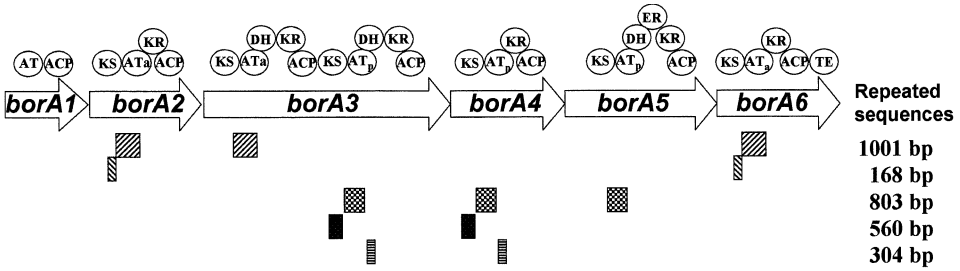

Figure 3. Scheme Representing the Repeated DNA Sequences within the Borrelidin PKS Genes

macrolactone prior to addition of a deoxyhexose moiety (41% identity, accession number AAD41818), SuaC from *S. griseolus* (39% identity; accession number AAA26823), and P-450sca-2 from *S. carbophilus*, involved in pravastatin biosynthesis (38% identity; accession number BAA06492). BorI is therefore proposed to catalyze oxidation of the C12-methyl group during borrelidin biosynthesis. The *borJ* gene resembles different PLP-dependent aminotransferases with the highest similarities with BioA from *Kurthia* sp. 538-KA26 (39% identity; accession number BAB39453), aminotransferase class III encoded by gene BA_4800 from *Bacillus anthracis* A2012 (38% identity; accession number NP_658160), and adenosylmethionine-8-amino-7-oxononanoate aminotransferase BioA from *Bacillus subtilis* (40% identity; accession number BAC03240). BorJ is therefore proposed to catalyze the introduction of a nitrogen atom into borrelidin at the activated C28-position, probably via a C12-formyl moiety. Mutants obtained by disruption of these genes failed to make any borrelidin, and experiments are underway in our laboratories to investigate the formation of the nitrile moiety.

Borrelidin Resistance

The *borO* gene product resembles different threonyl-tRNA synthetases, such as *ThrS* from *S. coelicolor* A3(2) (50% identity; accession number NP_625810). It has been shown that threonyl-tRNA synthetase is a molecular target for borrelidin in several bacterial strains [11], and, most probably, self-resistance of *S. parvulus* to borrelidin is due to the product of *borO*, which would encode a borrelidin-resistant threonyl-tRNA synthetase. Expression of *borO* in *S. albus* J1074 confers resistance to this organism to borrelidin (data not shown).

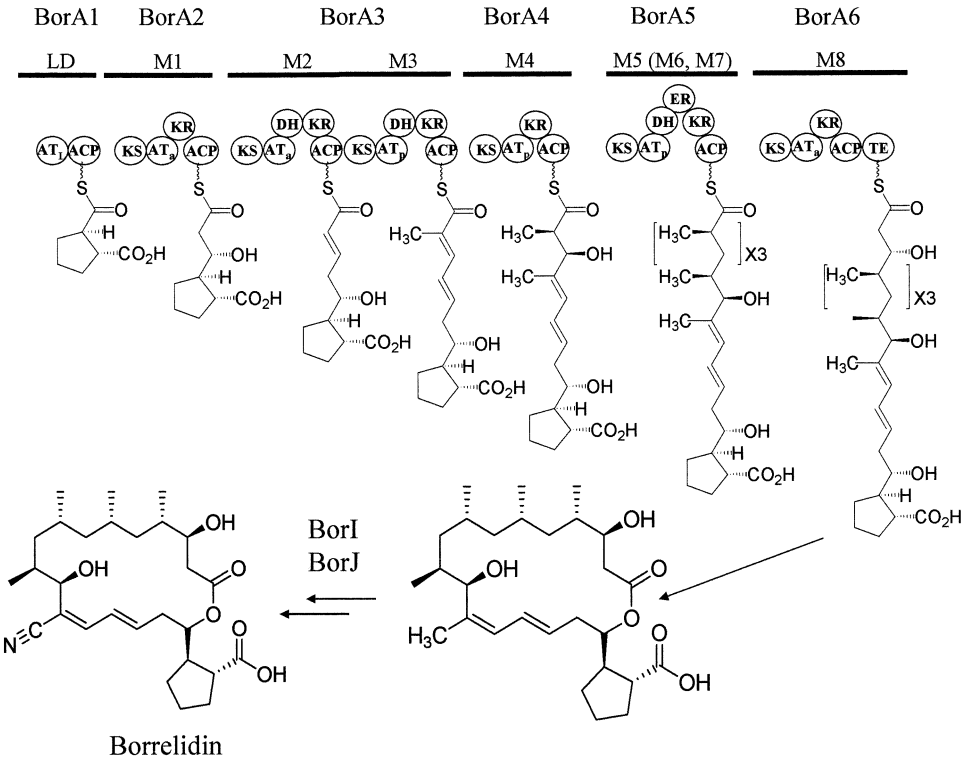

Figure 4. Module Organization in the Borrelidin PKS and Proposed Model for Borrelidin Biosynthesis

ACP, acyl carrier protein; AT, acyltransferase; AT_a, malonyl-specific acyltransferase; AT_p, methylmalonyl-specific acyltransferase; ER, enoyl reductase; KR, β-keto reductase; DH, dehydratase; KS, β-ketoacyl-acylcarrier protein synthase; TE, thioesterase.

Table 2. Production of Borrelidin by Different Mutants in the Borrelidin Gene Cluster

Gene	Plasmid	Mutant strain	Borrelidin Production[a]	
			No Addition	Plus *trans*-1,2-CPDA
wild-type	–	–	100 ± 16	363 ± 65
borC	pSLCr1	SPMC	0	933 ± 42
borD	pSLDr1	SPMD	0	75 ± 15
borE	pOJEd1	SPME	2 ± 1	122 ± 23
borF	pLHFr1	SPMF	3 ± 2	201 ± 52
borG	pOJGd1	SPMG	11 ± 1	1532 ± 142
borH	pOJHd2	SPMH-2	17 ± 2	203 ± 40
borL	pSLLr1	SPML	0	408 ± 70
borK	pSLKr1	SPMK	6 ± 1	468 ± 18
borM	pSLMr1	SPMM	0	461 ± 29
borN	pSLNr1	SPMN	25 ± 9	68 ± 12

[a] % relative to wild-type *S. parvulus* Tü4055 levels of borrelidin production.

Discussion

It is apparent from a retrobiosynthetic analysis that borrelidin biosynthesis includes several novel features, including the use of a dicarboxylic acid for the starter acid of a PKS and the incorporation of a nitrile moiety into a macrolide aglycone, either during polyketide assembly or as a post-PKS modification. The biosynthetic gene cluster responsible for borrelidin biosynthesis has been cloned and sequenced from the producing strain *S. parvulus* Tü4055 and has uncovered fascinating aspects of borrelidin biosynthesis at the molecular biology level. A total of 20 genes (spanning 52,369 bp) have been identified that have a role in borrelidin biosynthesis. All of these genes, with the exception of some PKS genes, have been inactivated through insertional mutation and subsequently complemented in *trans* to preclude the presence of polar effects when studying the phenotypes of gene inactivated strains. Based on sequence similarities, insertional inactivation, and feeding experiments, these genes fall into four broad categories: those involved in biosynthesis of the borrelidin polyketide backbone (i.e., the *bor* PKS), those involved in starter acid biosynthesis, those involved in formation of the nitrile moiety, and finally, those involved in regulation or resistance.

The *bor* PKS constitutes six polypeptides containing a total of seven modules, and only BorA3, comprised of extension modules 2 and 3, is bimodular. All the remaining modules, including the loading module, exist as discrete proteins (Figure 3). More commonly, the loading module is found together with the first extension module on one protein. The streptomycete-derived polyene macrolides nystatin [26] and amphotericin [27] and the myxobacterial epothilone [28] PKSs also exhibit this feature with *bor*. Many different starter units have been reported to be incorporated by loading domains of PKS [29]. The discrete *bor* PKS loading module selects and presents *trans*-cyclopentane-(1*R*,2*R*)-dicarboxylic acid (*trans*-1,2-CPDA) to the first extension module and represents the first loading module to be sequenced that naturally selects a dicarboxylic acid. The coronatine PKS incorporates a five-membered alicyclic monocarboxylic acid, namely 2-carboxycyclopent-2-eneone [30]. However, *borA1* is more similar to the loading modules of the avermectin and erythromycin PKSs. These have been shown to accept a number of alternative starter acids [31, 32], and notably, the avermectin loading module can accept cyclopentane monocarboxylic acid and cyclopent-1-ene monocarboxylic acid [31]. The AT domains of PKSs select a particular carboxylic acid unit as substrate. For extension modules, this selectivity has been shown to correlate with certain signature motifs within AT domains [18, 19]. These motifs are not conserved in the AT of the *bor* loading module, which is not surprising as this AT domain is the first to be sequenced that selects an alicyclic dicarboxylic acid. A further unusual aspect of the AT domain of the *bor* loading module is that the active site contains a GXCXG motif rather than the usual GXSXG. Substitution of the nucleophilic active site serine residue by cysteine is also known for two NRPS thioesterase domains, specifically those of the synthetases responsible for the production of mycobactin and pyochelin [33]. A GXCXG motif is also observed in a thioesterase-like domain of ORF1 in the bialaphos cluster [34]. Since it is not possible to interconvert between the two types of serine codon via a single base change, it has been suggested that active sites containing an essential serine residue lie on two lines of descent from an ancient ancestral enzyme that had a cysteine instead of a serine in its active site [35]. The presence of enzymes containing cysteine in the active site may support this view. It may alternatively be the case that cysteine arises in these active sites because it is possible to mutate from one type of serine codon to the other via a cysteine, which would remain catalytically active.

The identification of catalytic domains and determination of their boundaries in the *bor* PKS were based on similarities to the conserved amino acid sequences of other modular type I PKSs. All of the individual enzymatic domains appear to be catalytically active, based both on their amino acid sequences and on the amount of reductive processing required for each module based on a retrobiosynthetic analysis. This situation is rare, as modular type-I PKSs tend to include one or more inactive domains. A precedent for this is the spinosyn PKS, which likewise contains no inactive domains [36].

A particularly striking aspect of the *bor* PKS is the absence of two extension modules that would be required by the "one module, one extension" colinearity paradigm. Borrelidin contains a hydrophobic 1,3,5-tri-

methylhexyl moiety between positions C4 to C9 that requires three consecutive chain extensions, utilizing methylmalonyl-CoA as substrate, together with complete reduction (i.e., involving KR, DH, and ER domains) of the β-keto moiety generated during each chain extension cycle. Only BorA5 contains all three reductive domains, and furthermore the AT of this module selects for methylmalonyl-CoA. It could be possible that the *bor* PKS utilizes module 5 through three distinct chain extension cycles on one nascent acyl chain. If this is indeed the case, then to the best of our knowledge this represents only the third example in which a type I modular PKS contains a module that appears to be used iteratively (i.e., as opposed to aberrant iterative use of one module, referred to as "stuttering" [37, 38]), the other cases being the PKSs responsible for stigmatellin [39] and lankacidin [40] biosynthesis. However, in all of these cases, this conjecture is only based on comparison of molecular structure with the respective biosynthetic gene cluster, and no functional evidence has yet been presented. Recently, we have shown that engineered gene fusions of the *bor* PKS genes in which module 5 was translationally fused to either or both flanking modules (module 4 encoded by *borA4* and module 8 encoded by *borA6*) remained able to synthesize borrelidin [41]. These experiments indicate that it is unlikely that PKS modules encoded elsewhere on the genome are involved in borrelidin biosynthesis. Additionally, it appears improbable, based on these data, that three separate copies of BorA5 dock together and catalyze three independent rounds of chain extension in a processive manner. BorA5 therefore operates iteratively, although the specific mechanism governing its functionality remains to be uncovered [41].

The *bor* extension modules 1, 2, and 8 all incorporate malonyl-CoA and, strikingly, feature near total identity of the AT domains both at the amino acid and nucleotide level. The AT domains of modules 4, 5, and 8, which select for methylmalonyl-CoA, likewise display a very high level of sequence identity. The presence of these repeated regions may explain the deletions frequently observed while genetically manipulating *S. parvulus* Tü4055 (C.O., unpublished results).

The starter acid for borrelidin biosynthesis is *trans*-cyclopentane-1,2-dicarboxylic acid, as confirmed by feeding the commercially available racemate of this acid. Based on the absolute stereochemistry of borrelidin [6,7], *trans*-cyclopentane-(1*R*,2*R*)-dicarboxylic acid is the starter unit accepted by the *bor* loading module. Furthermore, racemic *cis*-cyclopentane-1,2-dicarboxylic acid, fed as the monoethylester (the *trans* equivalent of which is an excellent substrate), did not contribute to borrelidin biosynthesis in strains deficient in BorE (unpublished results). The biosynthetic pathway to the novel dicarboxylic acid starter unit *trans*-1,2-CPDA is complex. It has been resolved in part through gene inactivation with subsequent complementation in *trans* of the full-length gene and through feeding with exogenous *trans*-1,2-CPDA to these mutants. A number of genes were identified to be essential for *trans*-1,2-CPDA production: *borC*, *borD*, *borE*, *borF*, *borK*, *borL*, and *borM*. A further set of genes were found that may well have a role in *trans*-1,2-CPDA biosynthesis: *borG*, *borH*, and *borN* (see Table 2 and Results). Putative functions have been assigned to these genes (see Table 1 for similarities). Of particular interest is *borE*, a member of the enolase superfamily and therefore potentially able to stabilize the formation of a carbanion on the carbon adjacent to a carboxylate. It is reasonable to suggest that such a step may be involved in the formation of the cyclopentane ring by generating an α-carboxyl carbanion nucleophile (or equivalent) that could undergo intermolecular attack at an activated carbon at C2 (activated as either a carbonyl or phosphorylated secondary alcohol) (see Figure 5).

Hypothetically, the biosynthesis of *trans*-1,2-CPDA may originate from tyrosine catabolism. The product of *borN* exhibits homology to 2-hydroxyhept-2,4-diene-1,7-dioate isomerase, a key enzyme in 4-hydroxyphenylacetic acid catabolism, which is itself a product of tyrosine catabolism. BorN may catalyze the isomerase reaction for which the homolog has been characterized, i.e., the production of 2-oxohept-3-ene-1,7-dioate (Figure 5), with the conjugated double bond then being reduced by BorK or BorM. Either BorC or BorD could reduce the keto group at C2, and the resultant hydroxyl could be phosphorylated by the action of BorF. 2-phosphoheptane-1,7-dioate could then be cyclized (as described above) by the action of BorE to yield *trans*-1,2-CPDA. This scheme would then require thioesterification by coenzyme A in order to generate the activated substrate for the *bor* loading module. Since BorN would catalyze a reaction that is essentially one of primary metabolism, its function may be to channel substrate into *trans*-1,2-CPDA biosynthesis, thus forming a novel link between primary and secondary metabolism. When *borN* is inactivated, borrelidin, and hence *trans*-1,2-CPDA, is still biosynthesised, albeit at a reduced rate. This could be explained if 2-oxohept-3-ene-1,7-dioate produced via primary metabolism is utilized during *trans*-1,2-CPDA biosynthesis.

The hypothesis for *trans*-1,2-CPDA biosynthesis presented above is consistent with the genes shown to have a role in *trans*-1,2-CPDA biosynthesis, although clearly other biosynthetic routes are possible. The biosynthesis of *trans*-1,2-CPDA is currently being studied in our laboratories through precursor feeding and heterologous gene expression studies.

Feeding exogenous *trans*-1,2-CPDA (1 mM) to wild-type *S. parvulus* Tü4055 increased the yield of borrelidin 3.6 ± 0.6 fold, implying that starter acid biosynthesis is a bottleneck in borrelidin biosynthesis. Feeding *trans*-1,2-CPDA to all the mutants disrupted in one of the genes putatively involved in *trans*-1,2-CPDA biosynthesis (i.e., *borC*, *borD*, *borE*, *borF*, *borG*, *borH*, *borK*, *borL*, *borM*, or *borN*; see Table 2) reestablished borrelidin production, but to greatly different levels. No strong sequence homologies were found for *borG*, although it has a profound effect on borrelidin biosynthesis. The mutant containing a disrupted copy of *borG* loses most, though not all, of its ability to produce borrelidin. However, feeding this strain with exogenous starter acid increases the yield relative to wild-type 15.3 ± 1.3 fold (see Table 2). No functional or structural role could be assigned to *borG* in *trans*-1,2-CPDA production, but the 15-fold increase in borrelidin titer described above sug-

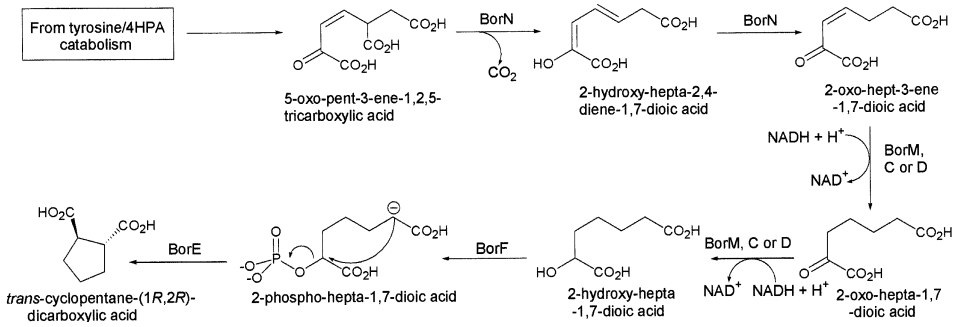

Figure 5. Putative Biosynthetic Pathway for Trans-1,2-CPDA from 4HPA

gests that *borG* may be involved in the regulation of *trans*-1,2-CPDA biosynthesis. A second gene likely to regulate *trans*-1,2-CPDA biosynthesis is *borL*, whose disruption leads to a borrelidin nonproducer. Although showing no strong resemblance to known bacterial proteins, it does demonstrate weak homology with the putative regulatory bacterial proteins from the IclR and GntR families in *S. coelicolor* A3(2) (accession numbers NP_628363 and NP_624887). The strongest match for BorL corresponds to auxin-responsive proteins from plants, specifically *Arabidopsis thaliana*. Auxins are plant hormones, and to the best of our knowledge auxins and auxin-response proteins have not been observed in bacteria.

The work presented here has revealed a number of unusual and interesting aspects of borrelidin biosynthesis and has highlighted remaining issues that will require further study through a combination of chemical, biochemical, and genetic analysis. These include details of *trans*-1,2-CPDA biosynthesis, formation of the nitrile moiety, and the apparently iterative use of a type I modular PKS module. It has also set the stage for a drug discovery program in which novel analogs of borrelidin can now be generated through genetic engineering of the biosynthetic genes.

Significance

Macrocyclic polyketides include some of the most important pharmaceutical agents in current clinical use. Although discovered because of its antibiotic activity, borrelidin, a nonglycosylated 18-membered macrolide, has been recently "rediscovered" due to its anti-angiogenesis activity and ability to induce apoptosis of capillary-forming cells, and therefore it has (or its derivatives have) great potential as future cancer treatments. Borrelidin also possesses several unique chemical features, including a nitrile moiety at C12 and an unusual starter unit of the polyketide chain at C17 (a *trans*-1,2-disubstituted cyclopentane dicarboxylic acid). The characterization of the borrelidin biosynthetic gene cluster isolated from *Streptomyces parvulus* Tü4055 has shown that its macrolide ring is formed by a modular polyketide synthase. Interestingly, only seven rather than the nine modules expected for a nonaketide product are encoded by the *bor* PKS. In addition, a set of genes were shown by sequence similarities, insertional inactivation, and feeding experiments to be involved in the biosynthesis of the starter unit. The identification of the borrelidin cluster opens up a way for future studies aimed at engineering more potent analogs of borrelidin as angiogenesis inhibitors and antitumor agents through combinatorial biosynthesis.

Experimental Procedures

Strains, Culture Conditions, and Plasmids

S. parvulus Tü4055 and mutants were routinely grown on tryptone soya broth (TSB). Production, transformation, and regeneration of protoplasts were performed following standard procedures [42]. *Streptomyces albus* J1074 [42] was used as host for expression of the borrelidin-resistance gene. When antibiotic selection of transformants was required, 100 μg/ml ampicillin, 20 μg/ml tobramycin, 25 μg/ml apramycin, and 50 μg/ml hygromycin were used when necessary.

DNA Manipulation

DNA manipulations were performed according to standard procedures for *E. coli* [43] and *Streptomyces* [42]. A cosmid library of *S. parvulus* Tü4055 genomic DNA was constructed. DNA fragments obtained from a partial digestion with Sau3AI were ligated to cosmid pWE15 digested with BamHI and in vitro packaged using the Gigapack III Gold packaging Extract kit according to the manufacturer's handbook (Stratagene). A number of the resulting *E. coli* transductants (3300 colonies) were picked and transferred to 96-well microtitre plates containing Luria broth (LB) medium and 100 μg/ml ampicillin. Clones were replica plated onto Luria agar (LA) plates containing ampicillin. After overnight growth at 37°C, colonies were transferred to nylon membrane filters for in situ colony hybridization analysis according to published methods [43] and screened using a labeled probe that was generated using the DIG DNA labeling and detection kit (Roche).

Sequencing of the Borrelidin Gene Cluster

Cosmids Bor32A2 and Bor19B9 were sequenced by shotgun sequencing of a Sau3AI-derived subclone library for each cosmid, consisting of 1.5–2.0 kb fragments in pHSG397. DNA sequencing was carried out using an Applied Biosystems 800 Molecular Biology CATALYST robot to perform the dideoxy termination reactions, which were then loaded into an ABI Prism 3700 automated sequencer (Applied Biosystems). The raw sequence data were processed using the Staden software package. Assembly and contig editing was performed using GAP (Genome Assembly Program) version 4.2 [44].

Generation of Mutants in *S. parvulus* Tü4055

To generate mutants in different genes of the borrelidin cluster, several plasmids were constructed for gene replacement or gene disruption experiments. Primers used for PCR amplification of the different genes and their flanking restriction sites are listed in the

Supplemental Data available with this article online. Cosmids Bor32A2 and Bor19B9 were used as DNA templates for subcloning or PCR amplification.

All final constructs were introduced into *S. parvulus* Tü4055 by protoplast transformation, and transformants were selected with apramycin. Resistant colonies were transferred to MA agar plates containing apramycin. In the case of gene replacement experiments, transformants were also grown without the presence of hygromycin, and apramycin-resistant hygromycin-sensitive colonies were selected for further analysis.

Complementation of *S. parvulus* Tü4055 Mutant Strains

To complement the mutations in the different mutant strains, several plasmid constructs were generated (see Supplemental Data), introduced into the appropriate *S. parvulus* Tü4055 mutant strains by protoplast transformation, and transformants were selected with hygromycin in R5 medium. Transformants were then transferred to MA agar plates containing apramycin and hygromycin, and after sporulation borrelidin production was assayed.

Feeding Experiments

Feeding experiments with *trans*-D,L-cyclopentane-1,2-dicarboxylic acid (Sigma) were performed in R5A [45] solid media (for routine testing) and PYDG liquid medium (for quantitative determinations). PYDG contains (g/l): peptonized milk nutrient, 15; yeast autolysate, 1.5; dextrin, 45; glucose, 5; adjusted with 5 M NaOH (aq.) to pH 7.0. The precursor (at a final concentration of 1 mM) was added at the beginning of incubation to the solid medium and after 24 hr of incubation when growing in liquid medium.

Analysis of Borrelidin Production

Borrelidin production was assessed qualitatively by growing *S. parvulus* wild-type strain or mutants on solid R5A medium. After 5 days at 30°C, agar plugs containing 1.5 ml of agar media were extracted with 1 ml ethyl acetate, and the presence of borrelidin in the extract was analyzed by HPLC as described below. Borrelidin production was assessed quantitatively in liquid medium. Strains were grown as follows: a seed culture containing NGY media (30 ml in a 250 ml Erlenmeyer flask) was inoculated with 0.5 ml of vegetative mycelia. After 2 days incubation in a rotary incubator (30°C, 250 rpm), 3 ml of the cultures was used to inoculate (10%, v/v) 30 ml PYDG production medium. After 5 additional days of incubation with shaking, the cultures were harvested for analysis. One milliliter aliquots of fermentation broths were removed and adjusted to pH ~3 by the addition of 90% formic acid (~50 μl). Ethyl acetate (1 ml) was added to the sample, and this was then mixed vigorously for 30 min. The phases were separated by centrifugation in a microfuge, and the upper phase was transferred to a clean 2 ml Eppendorf tube. The ethyl acetate was removed by evaporation using a Speed-Vac. Residues were resuspended in methanol (250 μl) and clarified by centrifugation. Analysis of borrelidin production was performed using the following HPLC system: injection volume, 50 μl; column stationary phase, 150 × 4.6 mm column, base-deactivated reversed phase silica gel 3 μm (Hypersil C_{18}-BDS); mobile phase A: 10% acetonitrile: 90% water, containing 10 mM ammonium acetate and 0.1% TFA; mobile phase B: 90% acetonitrile: 10% water, containing 10 mM ammonium acetate and 0.1% TFA; mobile phase gradient: T = 0 min, 25% B; T = 15, 100% B; T = 19, 100% B; flow rate, 1 ml/min; detection, UV at 258 nm (DAD acquisition over 190–600 nm).

Supplemental Data

Information regarding vectors, PCR conditions, generation of mutants in *S. parvulus* Tü4055, and complementation of mutants is available as at http://www.chembiol.com/cgi/content/full/11/1/87/DC1.

Acknowledgments

This research was supported by grants from the Spanish Ministery of Science and Technology (BIO2000-0274 and BMC2003-0478) and the Plan Regional de I+D+I del Principado de Asturias (GE-MED01-05). We thank Obra Social Cajastur for financial support to C.O. and C.S. The authors wish to thank Professor Wolfgang Wohlleben for providing us with the *Streptomyces parvulus* strain and Dr Antonio Jiménez for providing us with the *hyg* gene. We also thank Dr Sabine Gaisser for a number of important observations and suggestions and for critical reading of the manuscript and Dr. Christian B. Stark for synthesis of cis-cyclopentane-1,2-dicarboxylic acid.

Received: September 26, 2003
Revised: October 30, 2003
Accepted: October 30, 2003
Published: January 23, 2004

References

1. Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. Chem. Rev. *97*, 2465–2498.
2. Leadlay, P.F. (1997). Combinatorial approaches to polyketide biosynthesis. Curr. Opin. Chem. Biol. *1*, 162–188.
3. Staunton, J., and Wilkinson, B. (1997). Biosynthesis of erythromycin and rapamycin. Chem. Rev. *97*, 2611–2630.
4. Rawlings, B.L. (2001). Type I polyketide biosynthesis in bacteria. Nat. Prod. Rep. *18*, 231–281.
5. Keller-Scheirlein, W. (1967). Composition of the antibiotic borrelidin. Helv. Chim. Acta *50*, 731–753.
6. Kuo, M.S., Yurek, D.A., and Kloosterman, D.A. (1989). Assignment of 1H and 13C NMR signals and the alkene geometry at C-7 in borrelidin. J. Antibiot. (Tokyo) *42*, 1006–1007.
7. Anderson, B.F., Herlt, A.J., Richards, R.W., and Robertson, G.B. (1989). Crystal and molecular structures of two isomorphous solvates of the macrolide antibiotic borrelidin: absolute configuration determination by incorporation of a chiral solvent in the crystal lattice. Aust. J. Chem. *42*, 717–730.
8. Maehr, H., and Evans, R.H. (1987). Identity of borrelidin with treponemycin. J. Antibiot. (Tokyo) *40*, 1455–1456.
9. Berger, J., Jampolsky, L.M., and Goldberg, M. (1949). Borrelidin, a new antibiotic with anti-*borrelia* activity and penicillin enhancement properties. Arch. Biochem. *22*, 476–478.
10. Singh, S.K., Gurusiddaiah, S., and Whalen, J.W. (1985). Treponemycin, a nitrile antibiotic active against *Treponema hyodysenteriae*. Antimicrob. Agents Chemother. *27*, 239–245.
11. Paetz, W., and Nass, G. (1973). Biochemical and immunological characterization of threonyl-tRNA synthetase of two borrelidin-resistant mutants of *Escherichia coli* K12. Eur. J. Biochem. *35*, 331–337.
12. Dickinson, L., Griffiths, A.J., Mason, C.G., and Mills, R.F. (1965). Anti-viral activity of two antibiotics isolated from a species of *Streptomyces*. Nature *206*, 265–268.
13. Wakabayashi, T., Kageyama, R., Naruse, N., Tsukahara, N., Funahashi, Y., Kitoh, K., and Watanabe, Y. (1997). Borrelidin is an angiogenesis inhibitor; disruption of angiogenic capillary vessels in a rat aorta matrix culture model. J. Antibiot. (Tokyo) *50*, 671–676.
14. Matter, A. (2001). Tumor angiogenesis as a therapeutic target. Drug Discov. Today *6*, 1005–1024.
15. Tsuchiya, E., Yukawa, M., Miyakawa, T., Kimura, K.I., and Takahashi, H. (2001). Borrelidin inhibits a cyclin-dependent kinase (CDK), Cdc28/Cln2, of *Saccharomyces cerevisiae*. J. Antibiot. (Tokyo) *54*, 84–90.
16. Duffey, M.O., LeTiran, A., and Morken, J.P. (2003). Enantioselective total synthesis of borrelidin. J. Am. Chem. Soc. *125*, 1458–1459.
17. Swan, D.G., Rodríguez, A.M., Vilches, C., Méndez, C., and Salas, J.A. (1994). Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence. Mol. Gen. Genet. *242*, 358–362.
18. Reeves, C.D., Murli, S., Ashley, G.W., Piagentini, M., Hutchinson, C.R., and McDaniel, R. (2001). Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. Biochemistry *40*, 15464–15470.
19. Del Vecchio, F., Petkovic, H., Kendrew, S.G., Low, L., Wilkinson, B., Lill, R., Cortes, J., Rudd, B.A., Staunton, J., and Leadlay, P.F. (2003). Active-site residue, domain and module swaps in modular polyketide synthases. J. Ind. Microbiol. Biotechnol. *30*, 489–494.

20. Butler, A.R., Bate, N., and Cundliffe, E. (1999). Impact of thioesterase activity on tylosin biosynthesis in *Streptomyces fradiae*. Chem. Biol. *6*, 287–292.
21. Doi-Katayama, Y., Yoon, Y.J., Choi, C.Y., Yu, T.W., Floss, H.G., and Hutchinson, C.R. (2000). Thioesterases and the premature termination of polyketide chain elongation in rifamycin B biosynthesis by *Amycolatopsis mediterranei* S699. J. Antibiot. (Tokyo) *53*, 484–495.
22. Heathcote, M.L., Staunton, J., and Leadlay, P.F. (2000). Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units. Chem. Biol. *8*, 207–220.
23. Prieto, M.A., Díaz, E., and García, J.L. (1996). Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli* W: engineering a mobile aromatic degradative cluster. J. Bacteriol. *178*, 111–120.
24. Schmidt, D.M.Z., Hubbard, B.K., and Gerlt, J.A. (2001). Evolution of enzymatic activities in the enolase superfamily: functional assignment of unknown proteins in *Bacillus subtilis* and *Escherichia coli* as L-Ala-D/L-Glu epimerases. Biochemistry *40*, 15707–15715.
25. Walker, L., and Estelle, M. (1998). Molecular mechanisms of auxin action. Curr. Opin. Plant Biol. *1*, 434–439.
26. Brautaset, T., Sekurova, O.N., Sletta, H., Ellingsen, T.E., Strom, A.R., Valla, S., and Zotchev, S.B. (2000). Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. Chem. Biol. *7*, 395–403.
27. Caffrey, P., Lynch, S., Flood, E., Finnan, S., and Oliynyk, M. (2001). Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. Chem. Biol. *8*, 713–723.
28. Molnar, I., Schupp, T., Ono, M., Zirkle, R., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D.D., et al. (2000). The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. Chem. Biol. *7*, 97–109.
29. Moore, B.S., and Hertweck, C. (2002). Biosynthesis and attachment of novel bacterial polyketide synthase starter units. Nat. Prod. Rep. *19*, 70–99.
30. Rangaswamy, V., Jiralerspong, S., Parry, R., and Bender, C.L. (1998). Biosynthesis of the *Pseudomonas* polyketide coronafacic acid requires monofunctional and multifunctional polyketide synthase proteins. Proc. Natl. Acad. Sci. USA *95*, 15469–15474.
31. Pacey, M.S., Dirlam, J.P., Geldart, R.W., Leadlay, P.F., McArthur, H.A., McCormick, E.L., Monday, R.A., O'Connell, T.N., Staunton, J., and Winchester, T.J. (1998). Novel erythromycins from a recombinant *Saccharopolyspora erythraea* strain NRRL 2338 pIG1. I. Fermentation, isolation and biological activity. J. Antibiot. (Tokyo) *51*, 1029–1034.
32. Holzbaur, I.E., Ranganathan, A., Thomas, I.P., Kearney, D.J., Reather, J.A., Rudd, B.A., Staunton, J., and Leadlay, P.F. (2001). Molecular basis of Celmer's rules: role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. Chem. Biol. *8*, 329–340.
33. Shaw-Reid, C.A., Kelleher, N.L., Losey, H.C., Gehring, A.M., Berg, C., and Walsh, C.T. (1999). Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of *E. coli* EntF catalyzes both elongation and cyclolactonization. Chem. Biol. *6*, 385–400.
34. Raibaud, A., Zalacain, M., Holt, T.G., Tizard, R., and Thompson, C.J. (1991). Nucleotide sequence analysis reveals linked N-acetyl hydrolase, thioesterase, transport, and regulatory genes encoded by the bialaphos biosynthetic gene cluster of *Streptomyces hygroscopicus*. J. Bacteriol. *173*, 4454–4463.
35. Brenner, S. (1988). The molecular evolution of genes and proteins: a tale of two serines. Nature *334*, 528–530.
36. Waldron, C., Matsushima, P., Rosteck, P.R., Jr., Broughton, M.C., Turner, J., Madduri, K., Crawford, K.P., Merlo, D.J., and Baltz, R.H. (2001). Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*. Chem. Biol. *8*, 487–499.
37. Wilkinson, B., Foster, G., Rudd, B.A.M., Taylor, N.L., Blackaby, A.P., Sidebottom, P.J., Cooper, D.J., Dawson, M.J., Buss, A.D., Gaisser, S., et al. (2000). Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. Chem. Biol. *7*, 111–117.
38. Hardt, I.J., Steinmetz, H., Gerth, K., Sasse, F., Reichenbach, H., and Höfle, G. (2001). New natural epothilones from *Sorangium cellulosum*, strains So ce90/B2 and So ce90/D13: isolation, structure elucidation, and SAR studies. J. Nat. Prod. *64*, 847–856.
39. Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R. (2002). The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. J. Biol. Chem. *277*, 13082–13090.
40. Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K., and Kinashi, H. (2003). The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. Mol. Microbiol. *48*, 1501–1510.
41. Olano, C., Wilkinson, B., Moss, S.J., Braña, A.F., Méndez, C., Leadlay, P.F., and Salas, J.A. (2003). Evidence from engineered gene fusions for the repeated use of a module in a modular polyketide synthase. Chem. Commun., 2780–2782.
42. Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, C.A. (2000). Practical Streptomyces Genetics (Norwich, United Kingdom: The John Innes Foundation).
43. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory press).
44. Bonfield, J.K., Smith, K., and Staden, R. (1995). A new DNA sequence assembly program. Nucleic Acids Res. *23*, 4992–4999.
45. Fernández, E., Weissbach, U., Reillo, C.S., Braña, A.F., Méndez, C., Rohr, J., and Salas, J.A. (1998). Identification of two genes from *Streptomyces argillaceus* encoding glycosyltransferases involved in transfer of a disaccharide during biosynthesis of the antitumor drug mithramycin. J. Bacteriol. *180*, 4929–4937.

Accession Numbers

The complete DNA sequence of the borrelidin gene cluster has been deposited in EMBL under accession number AJ580915.